

# **Private Screening Operations Performance Evaluation Report**



## **Transportation Security Administration**

**Submitted By:**



**April 16, 2004  
Summary Report**

## TABLE OF CONTENTS

<b>1.0</b>	<b>EXECUTIVE SUMMARY .....</b>	<b>1</b>
1.1	Interpretation of Findings .....	1
1.2	Findings .....	2
1.3	Private and Federal Screening Model Strengths .....	3
<b>2.0</b>	<b>INTRODUCTION .....</b>	<b>6</b>
2.1	Evaluation Purpose .....	6
2.2	Statement of the Problem .....	6
2.3	TSA's Objectives for the Screening Pilot Program .....	7
2.4	Evaluation Questions .....	8
<b>3.0</b>	<b>BACKGROUND .....</b>	<b>9</b>
3.1	Private Pilot Program Overview .....	9
3.2	ATSA Interpretation and Degrees of Freedom .....	9
<b>4.0</b>	<b>EVALUATION PROCESS .....</b>	<b>11</b>
4.1	Evaluation Activities .....	11
4.2	Evaluation Design .....	12
4.3	Evaluation Methodology .....	13
<b>5.0</b>	<b>CONCLUSIONS .....</b>	<b>14</b>
5.1	Security Effectiveness Criteria .....	14
5.2	Customer Satisfaction and Stakeholder Impact Criteria .....	15
5.3	Cost .....	16
5.4	Uses and Limitations of the Study .....	18

### 1.0 EXECUTIVE SUMMARY

This report is a summary of the methodology and findings discussed in the “Private Screening Operations Performance Evaluation Report.” The report provides an evaluation of the performance of the five commercial airports that retained private screening personnel under Transportation Security Administration (TSA) oversight after the federalization of the nation’s screening workforce in 2002. The five airports are: San Francisco International, CA (SFO); Kansas City International, MO (MCI); Greater Rochester International, NY (ROC); Jackson Hole, WY (JAC); and Tupelo Regional, MS (TUP). The purpose of the evaluation was to determine whether the performance of private screening operations was equal to or greater than that provided by the federal government. This standard was established under the Aviation and Transportation Security Act (ATSA) for certification of private screening companies seeking to provide service at the nation’s airports.

TSA hired BearingPoint as an independent evaluator to conduct this study. BearingPoint, working with Abt Associates, developed three criteria to compare the performance of private screening operations to federal screening operations. Those criteria are: security effectiveness, cost, and customer and stakeholder impact. For each of the criteria, the Evaluation team developed quantitative measures to compare performance. The team also examined the private screening program to understand the qualitative factors that may impact performance. During the course of the Evaluation, the team made 29 airport visits, conducted over 240 interviews with federal and private sector personnel involved in screening operations, collected data related to screening operations, conducted statistical analysis, and developed the findings.

### 1.1 INTERPRETATION OF FINDINGS

The findings of this study must be viewed in light of five key factors:

- First, the study is best understood as a comparison of government delivery of screening services to delivery by a public/private hybrid. The private screening environment is not a pure privatization model. In the private screening operations, each of the three criteria above is influenced by decisions made by federal oversight and management. Therefore, the team was not able to discern in many cases whether the performance level is due to the private contractor or federal oversight staff at the privately screened airport.
- Second, the design of the private screening program severely limits the opportunity for differences in the two models. The private screening program was designed to ensure that factors driving security effectiveness and cost were similar in the two models. However, the private screening operations were granted some latitude in the execution of their responsibility to provide screening services, and the exercise of that latitude does reveal itself in the comparison.
- Third, the period in which the study was conducted provided several challenges. The program’s 15 month existence, from November 19, 2002 to present, was not a steady state environment in which to conduct the study. This period involved a start-up phase for both TSA and the private contractors and a significant workforce adjustment at the midpoint.
- Fourth, the airports selected for the program are few in number (5) and were chosen to provide variation in size, passenger type, and other characteristics. The TSA had little discretion in this area as the number of airports in the pilot program was limited by statute. The fact that these airports were not chosen at random and the small number of pilot airports seriously limits the

program's usefulness as a true scientific "pilot." Therefore, the program design limits the ability of the findings to be generalized to apply to future privately screened airports.

- Finally, the data available for review and analysis is limited. No historical baseline exists for the comparison. Because screening procedures, equipment, and screener compensation and training are vastly different today than they were in the pre-TSA environment, a useful baseline is not available for the five privately screened airports. In addition, TSA's financial management, human resources, and performance management systems were still evolving during the period studied.

The ability to generalize conclusions from this report and apply them to any future expansion of privately screened airports is limited by the factors cited above. Additional factors that may limit the number of differences in the findings are the "degrees of freedom" discussed in Section 3.2. TSA may want to address these issues in any future decisions regarding private screening operations. If TSA desires a more robust comparison of private screening operations to federal screening, it should consider three steps: open some of the degrees of freedom in a controlled manner; provide a larger, well-designed sample of airports; and improve its data collection systems.

## 1.2 FINDINGS

In the areas of security effectiveness and customer and stakeholder impact, the study compared the performance of each privately screened airport to a set of comparable federally screened airports. A determination was made as to whether the airport outperformed, under-performed, or performed at the same level as its federal counterparts. In the area of cost, the study compared the actual cost of screening at a privately screened airport to an estimate of what it would have cost the government if TSA had provided screening personnel at the same airport.

In general, the Evaluation team found that privately screened airports have met the standard articulated in ATSA that contract screening operations perform at the same level or better than federally screened operations to be certified to conduct screening under the statutorily authorized opt out program. The statistical analysis finds no evidence that they are not meeting the ATSA standard.

Findings in each of the three criteria areas are as follows:

- In the area of security effectiveness, there is no evidence that any of the five privately screened airports performed below the average level of the federal airports as measured by covert testing, Threat Image Projection (TIP) data, gate screening and recertification testing. However, there is credible evidence that Kansas City is outperforming the average level of its federal counterparts.
- In the area of cost to the government, costs for the five privately screened airports were not significantly different from the estimated cost of a federal screening operation at that same airport.
- In the area of customer satisfaction, performance of the privately screened airports compared to the federally screened airports was mixed in Categories X and I, and inconclusive in Categories II, III, and IV.<sup>1</sup> A qualitative survey of stakeholders revealed no difference between privately and federally screened airports.

---

<sup>1</sup> Sufficient customer satisfaction data was unavailable for Categories II, III, and IV.

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



Findings for each of the five privately screened airports are as follows:

- At San Francisco, a Category X airport, there is no evidence that it is different than federally screened airports within its category in security effectiveness. In the area of cost, the estimated cost of screening conducted by the federal workforce was not materially different than the cost of private screening at that airport. There were mixed results on customer satisfaction when compared to its federal counterparts. Passengers had less confidence in the security process but experienced shorter wait times.
- There is evidence that Kansas City, a Category I airport, is outperforming federally screened airports within its category in security effectiveness. In the area of cost, estimated costs of screening conducted by a federal workforce were higher than the costs experienced with contractors. However, these differences are just within the range of prediction error and the team is unable to conclude with certainty that the cost of federal screening would have been higher at this airport. There were mixed results on customer satisfaction when compared to its federal counterparts. Passengers experienced more thorough screening and shorter wait times, but less overall customer satisfaction.
- At Rochester International (ROC), a Category II airport, Jackson Hole (JAC), a Category III airport, and Tupelo Regional (TUP), a Category IV airport, the limited evidence available does not show differences in the area of security effectiveness. In the area of cost, the estimated cost of screening conducted by a federal workforce was not materially different than the cost of private screening at these airports. In the area of customer satisfaction, there was not enough data available to support any conclusion.

A detailed discussion of each finding is provided in section 5.0 of this document. The report also provides a description of the private screening program and a discussion of the approach, measures, analysis, and findings for each of the three criteria.

### 1.3 PRIVATE AND FEDERAL SCREENING MODEL STRENGTHS

The following section describes strengths of the private contractor and federal security screening models. These qualitative observations are a result of the field interviews conducted by the Evaluation team at the 29 airports visited. The team believes program strengths to be the result of numerous factors that affect screening operations. For example, the degrees of freedom provided by the Private Pilot Program, TSA's contract vehicle, and a hybrid management structure are all factors that contribute to the strengths of the private contractor and federal security screening models.

TSA authorized the private contractor screening companies some latitude in their implementation of security screening operations. This latitude is applicable to the infrastructure required to support the regulated security screening function. TSA termed this latitude the "degrees of freedom." As a result of variations in the contractors' implementation of these degrees of freedom, strengths in the federal and private contractor screening models can be observed.

The contract vehicle also contributes to differences in the two models. The private contractors have contractual obligations such as performance award fees, which the federally screened airports do not. The contract vehicle requires the private screening companies to uphold the same standard of security effectiveness as their federal counterparts, yet at the same time, it provides contractors a financial incentive to do so. TSA used cost plus contracting structures to allow for flexibility in an uncertain operating environment.

In addition, the hybrid management structure, where both the FSD (and staff) and the contractors play a role in the screening operations, distinguishes the private screening airports from their federal counterparts. The FSD and staff at privately screened airports are required to assume additional responsibilities and perform different roles than those required at federal screening airports. This too creates strengths in the two models. A list of each screening model's strengths is provided below.

#### 1.3.1 FEDERAL SCREENING MODEL STRENGTHS

- **Fewer Layers of Management** – Because there is no contractor interface between the FSD staff and the screener, the federal screening model has fewer layers of management thereby supporting faster and more efficient communication between management and the screener workforce.
- **Roles and Responsibilities are Clearly Defined** – The federal model provides a clear chain of command with little overlap of managerial responsibilities. However, under the private contractor model, certain roles and responsibilities are replicated within the FSD staff and the contract screening company. For example, both TSA and the contractor are providing supervision of screening personnel at the privately screened airports.
- **Direct Control of Screening Resources** – In the federal model, the FSD does not have to work through the contractor management team to address issues with the screener personnel. These issues include scheduling, staffing, employee behavior and attendance. In the federal model, the FSD has direct control of the screener workforce.
- **Communication with Stakeholders** – Under the federal model, the stakeholders know whom to address when communicating issues regarding screening operations. They know that the FSD has the authority to address issues at the screener level, and, as mentioned above, there are fewer layers of management to go through in addressing these issues. At some private contractor screening airports, stakeholders are unsure as to whom to talk to when addressing issues that arise at the screener level. They are unsure whether they should speak with the contractor or TSA staff.
- **Ability to Shift Resources within the Hub/Spoke System** – Certified federal screeners are authorized to perform screening duties at any federal airport regardless of location. Conversely, private contractor screeners, under the current policy, currently may not perform screening duties at federal airports. The current policy regarding the ability to deploy federal screener resources to different airports allows the FSD to address staffing shortages and unanticipated circumstances that affect security-screening operations.

#### 1.3.2 PRIVATE SCREENING MODEL STRENGTHS

- **Faster Decision-Making Process** – Because the private contractors have the ability to make managerial decisions at the local level, they can react more quickly and address issues in a more timely manner. In many cases, the federal screening operations must channel decisions through a centralized headquarters.
- **More Efficient Employee Discipline and Termination Process** – Private contractors are not required to comply with federal human resources administrative requirements, including government procedures related to the termination of employees. Contractors have the ability to terminate employees more expeditiously when needed. The time needed to terminate a sub-standard employee at a private contractor airport is significantly shorter than the time needed to

terminate a sub-standard employee at a federal airport. Due to the faster termination process under the private contractor-screening model, the contractor is better equipped to control absenteeism. Due to the private contractors' exemption from compliance with federal human resources administrative regulations, contractors are able to discipline screeners for poor performance in a more expeditious manner.

- **More Flexibility in Scheduling Screeners** – Private screening companies have the ability to schedule screening resources in a manner that best supports the screening operation. For example, the private contractors can split shifts as many times as is necessary to meet the peak passenger volume flow experienced at a particular airport. Conversely, under current policy the schedule for a federal screener can be split no more than once during the course a workday.
- **More Efficient Use of Personnel to Perform Non-Screening Functions** – Private contractor screening companies have the latitude to use less expensive resources to fulfill certain non-security related tasks. Due to a lack of TSA-authorized administrative support personnel at the federal airports, many FSDs have had to use screener personnel to perform administrative tasks such as HR and payroll support. In some cases, FSDs have used screeners that were placed on light duty to perform these tasks, while in other cases certified screeners are removed from the operational job they have been trained to perform. Private contractors use less expensive administrative staff to perform these functions. In addition to administrative tasks, contractors are using less expensive resources to perform other non-security related functions. For example, one private contractor has made use of non-screening personnel for baggage handling to save costs. This reduces the need to have trained screeners providing non-screening tasks.
- **Additional Compensation Opportunities** – Private contractors have the ability to provide screening personnel with additional compensation such as referral bonuses and other monetary rewards for performance. This creates additional incentives for screener personnel to achieve higher level of performance.
- **Cost Innovations** – Even though the contract vehicle does not explicitly encourage cost innovations, the private screening companies do have more of an incentive to reduce costs than their federal counterparts. The private screening companies would like to continue providing services to TSA beyond the interval stipulated by the Private Pilot Program. Therefore, the contractors must demonstrate they can operate in a cost-effective manner. Therefore, many contractors seek out ways in which they can innovatively reduce costs.
- **FSD Can Act as an Objective Observer when Dealing with Stakeholder Complaints about Screening Operations** – Because the screeners at a private contractor airport are not government employees, the FSD is able to take a more objective approach when dealing with screener-related issues raised by stakeholders such as airport management or air carriers.
- **More Visibility of Operational Performance** – Due to the award fee criteria stipulated in private contractor security screening contracts, TSA headquarters and field staff monitor the contractors' screening performance closely. As a result of this increased visibility, operational issues at the private pilot airports draw attention more easily.

Because the Private Pilot Program structure encourages the private screening companies to seek out and develop screening operations' efficiencies currently unavailable to federal screening operations, the airports participating in the Private Pilot Program exhibit more strengths in this arena than their federal counterparts.

## 2.0 INTRODUCTION

In November 2001, ATSA established the Transportation Security Administration to improve the security of all transportation modes within the United States. ATSA established several critical milestones with the intent of achieving a secure air travel system. The first milestone was to federalize the security screening operations at the nation's commercial airports by November 19, 2002. Section 108 of ATSA mandated that private security screening personnel be retained at up to five commercial airports under TSA oversight. As a result, TSA created five airport pilots. This 2-year pilot of private security screening operations under TSA oversight provides the opportunity to compare federal screening operations performance against that of the private sector.

Starting in November 2004, ATSA allows airports to petition TSA to opt-out of federalized security screening operations and to return to the use of private contractor screeners under certain parameters articulated in the statute. A key criterion for determining whether to grant such petitions is whether the level of screening services provided by the private contract screening operations is equal to or greater than that of federal screening operations. The ability to evaluate and measure the performance of both private contract and federal screening operations accurately will have a critical impact on the future security of air travel.

## 2.1 EVALUATION PURPOSE

ATSA charged TSA with responsibility for evaluating federal and private security screening operations performance, without specifying metrics or methods to be used in a performance evaluation. ATSA also charged TSA with reviewing airports' opt-out petitions and assigned TSA the responsibility to administer petitions, without specifying certification processes or petition evaluation criteria for approval or denial.

To evaluate the performance of federal and private security operations, TSA hired BearingPoint, Inc., as an independent evaluator. BearingPoint, working with Abt Associates, Inc., developed an impartial process for analyzing current private contractor and federal screening operations performance. The results of this evaluation will inform TSA with regard to policy decisions to be made in the definition of the opt-out program. It will also assist TSA in the development of the process to review opt-out petitions.

## 2.2 STATEMENT OF THE PROBLEM

In the wake of the terrorist attacks on September 11<sup>th</sup>, 2001, Congress sought to improve aviation security by overhauling existing procedures for screening passengers and their baggage. Investigations by the General Accounting Office (GAO) and others found in repeated instances of inadequate protection against contraband, highlighted the fact that no national system of certification or standard security procedures existed, and criticized the way in which screeners were recruited and deployed.<sup>2</sup> Employee background checks were inadequate; training was minimal, as were wages; and screener turnover rates were high—as high as 100% in some cities. This resulted in substantial part from the structural arrangements that were relied upon for aviation security prior to 9/11. Authority for safety and security was fragmented. Airport operators (typically public entities) were responsible for airport safety and were

---

<sup>2</sup> U.S. General Accounting Office, *Aviation Security: Long Standing Problems Impair Airport Screeners' Performance*, GAO/RCED-00-75 (Washington, D.C.: June 28, 2000); *Aviation Security: Vulnerabilities Still Exist in the Aviation Security System*, GAO/T-RCED/AMID-00-142 (Washington, D.C.: April 6, 2000).

regulated in this domain by the Federal Aviation Administration (FAA). Authority for airline security rested with custodial airlines, which generally hired private security firms to conduct passenger and baggage screening, and again regulated by the FAA. Responsibility and accountability was thereby split among a multitude of different carriers and airport operators. The financial incentives also encouraged security screening at low cost, especially given that the American public did not see domestic air piracy as a major national security threat.

Seeking to establish new laws and procedures to improve aviation security, the House and the Senate developed differing approaches. The compromised bill and the resulting law, ATSA, largely adopted the Senate's preference for federalization, but authorized a pilot program whereby passenger and baggage screening at five selected airports was to be performed by qualified private contractors (Public Law 107-71, Sec. 44919). In addition, at the end of two years following the law's passage, airport operators could petition the Administrator of TSA to "opt out" of the federalized program and have TSA manage screening services locally by means of TSA contracting with qualified private security firms under TSA oversight.

### **2.3 TSA'S OBJECTIVES FOR THE SCREENING PILOT PROGRAM**

Both the pilot program and the federalized screening operation seek the same goals—to increase aviation security by upgrading the quality of screener workforce, screen all baggage and all passengers effectively, and create an atmosphere that aligns the needs for security with the public's freedom of movement. Consistent with the Government Performance and Results Act of 1993 (GPRA), ATSA directs TSA to develop a plan for measuring the agency's effectiveness. In keeping with the goals of GPRA, these measures need to be concrete and results-oriented to assist TSA in making strategic policy decisions in the future. During the first six months, TSA was required to "(A) establish acceptable levels of performance for aviation security, including screening operations and access control, and (B) provide Congress with an action plan, containing measurable goals and milestones, that outlines how those levels of performance will be achieved." Over the longer term, the "the Secretary and the Under Secretary for Transportation Security shall agree on a performance plan for the succeeding five years that establishes measurable goals and objectives for aviation security. The plan shall identify action steps necessary to achieve such goals" (Section 130). Although the agency has established the performance plan as required by law, GAO observes that "to date these tools have focused on TSA's progress in meeting deadlines to implement programs and initiatives mandated by ATSA, rather than on the effectiveness of these programs and initiatives. TSA has recognized that its data on the effectiveness of its aviation security initiatives are limited and is taking steps to collect objective data to assess its performance, which is to be incorporated in the Department of Homeland Security's (DHS) 5-year performance plan."<sup>3</sup>

Congress created the pilot test—and the subsequent "opt-out" program—to leverage the private sector's creativity and flexibility to provide screening services potentially more efficiently than would be possible in a strictly federal model, while maintaining federal oversight. Absent established measures that TSA has defined for its own evaluative purposes, the present study assesses the effectiveness of screening performance, efficiency of screening operations, and public satisfaction with screening services in privately screened airports as compared to federally screened airports.

---

<sup>3</sup> General Accounting Office, *Aviation Security: Efforts to Measure Effectiveness and Address Challenges*. [GAO-04-232T](#) (November 5, 2003), p. 2.

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



### 2.4 EVALUATION QUESTIONS

The purpose of the Evaluation is to present a sound and unbiased performance evaluation comparing the efficiency and effectiveness of federal and private screening operations. The Evaluation was designed to answer three fundamental questions:

- How does each private pilot airport perform relative to other federal airports in that same category?
- What is the difference between what it costs the government to have private contractors perform screening services and what it costs the government to have federal screening at each of the private pilot airports for the same time period?
- What are the satisfaction levels of passengers and stakeholders at private pilot airports relative to the satisfaction levels of passengers and stakeholders at federally screened airports in that same category?

### **3.0 BACKGROUND**

This section provides background information regarding the Private Pilot Program. It describes the program's establishment and structure, the latitudes afforded to the contractors in conducting the screening operations, and the individual characteristics of each of the pilot airports.

#### **3.1 PRIVATE PILOT PROGRAM OVERVIEW**

ATSA set up a Private Pilot Program that would allow TSA to assess the level of screening performance achieved by private screening companies and determine how it compares to the level of performance achieved at federally screened airports. ATSA provided the following direction in establishing the program:

- The program's duration will be two years, beginning November 19, 2002,
- Airports must submit an application in order to participate in the program,
- Participation in the program is limited to one airport within each of the five airport security risk categories as defined by TSA,
- Federal supervisors will oversee the security screening operations,
- The private screening company is required to be owned and controlled by a citizen of the United States,
- TSA may terminate the contract entered into with a private screening company if that company fails to comply with any standard regulation, directive, order, law, or contract applicable to the hiring or training of personnel or to the provision of screening at the airport,
- Private screening companies must provide screeners comparable compensation and benefits as the federal screeners,
- Employees of the private screening companies must meet the same job-related standards as federal screeners, and
- At the end of the 2-year pilot period, the airport operator may elect to continue to employ private contract screening.

It is important to note that ATSA was not specific in defining the program structure and constraints. Therefore, TSA constructed a pilot program based on its interpretation of the law.

#### **3.2 ATSA INTERPRETATION AND DEGREES OF FREEDOM**

ATSA states that federal supervisors will oversee the security screening operations performed by the private screening companies. It does not, however, state where the role of the federal oversight function ends and the role of the screening contractor begins. As a result, TSA allows private screening companies to exercise limited discretion in the provision of some screening support functions. TSA calls these areas of discretion "degrees of freedom." Despite these degrees of freedom, the private contractors are held to the same overall security regulations as the federal security screeners, as defined in the TSA Standard Operating Procedures (SOP).

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



The private screening companies are permitted latitude by TSA in the following degrees of freedom:

- **Recruiting of screening personnel:** Private screening companies are allowed latitude in their methods of recruiting personnel in preparation for the TSA assessment process. Contractor personnel must pass the TSA assessment process and are required to meet all qualifications applicable to federal security screeners.
- **Supervision and scheduling of screening personnel:** Private screening companies have the latitude to staff at or below TSA-specified full-time equivalent (FTE) levels as long as security screening performance regulations (e.g., checkpoint manning levels as stated in the SOP) are met. Private screening companies are not required to staff with a specified number of full-time or part-time employees and can use various staffing models (e.g., all part-time, split shifts, management of down time) to address scheduled and unanticipated passenger and checked baggage volumes. Private screening companies employ their own screening managers, allowing TSA Screening Managers at the private contractor airports to perform a somewhat different function than they perform at airports with TSA screeners.
- **Management of materials and overhead:** Private screening companies have latitude in the negotiation of real estate, utilities, and other overhead costs associated with support of their security screening personnel.

Additional areas in the screening operation require strict adherence to TSA-mandated requirements. TSA decided that these functions must be the same at all airports to ensure the same security standard and level of screener qualification between federally screened and privately screened sites. Areas in which private screening companies do not have latitude and cannot deviate from TSA federal regulations are:

- **Screener assessment:** Private screening companies are required to comply with TSA criteria used to conduct assessments. Under the current standard, all private contractor screeners must undergo assessments performed by a national TSA-approved contractor. However, the program office has solicited input from the private screening contractors to perform assessments locally.
- **Screener training:** Private screening companies are required to receive the same basic training curriculum used by TSA to certify and re-certify security screening personnel. Private contractor screeners are also required to receive all recurring basic training that relates to security screening operations.
- **Staffing ceilings:** Private screening companies are prohibited from staffing above the ceilings established by TSA.
- **Screening SOPs:** Private screening companies are prohibited from any variance or deviation from TSA security screening SOPs.
- **Screener compensation:** Private screening companies must provide comparable compensation and benefits to that paid to federal screeners. (This degree of freedom was specified in ATSA.)
- **Screening equipment:** Private contractor screeners are required to operate and maintain security screening equipment at the same levels of performance as the federal security screeners.

### 4.0 EVALUATION PROCESS

In order to conduct the Screening Operations Performance Evaluation, the team divided the project into multiple phases. Phase I served as the investigation and planning phase and was completed in November 2003. During this phase, the team developed a model, which it would use to evaluate the performance of federal and private contractor security screening operations. A detailed description of Phase I activities is found in the Screening Operations Performance Evaluation Plan (Phase I Report). Phase II served as the data collection and analysis phase. During this phase, the team implemented the evaluation plan that included the collection and analysis of data relevant to the comparison of screening operations performance. This report documents the results of the Evaluation performed in Phase II. The team also identified program improvement initiatives applicable to both the Private Pilot Program and all airports in a separate report – Screening Operations Report: Program Improvements.

The timeline for each phase of the Evaluation is:

- Phase I: October 1, 2003 – November 14, 2003.
- Phase II: December 1, 2003 – April 9, 2004.

### 4.1 EVALUATION ACTIVITIES

#### Phase I

In Phase I, the team worked with TSA and external stakeholders to develop an independent and statistically sound evaluation model and plan. This phase consisted of the following activities:

- Develop a project plan for the Evaluation,
- Review existing data and identify data sources that would add value to the Evaluation,
- Observe screening processes and conduct interviews at the five private pilot airports and four federal airports,
- Solicit input from external stakeholders by convening a focus group,
- Define metrics for screening performance analysis,
- Scientifically select federal comparison airports for additional data collection, and
- Develop a process for data analysis.

#### Phase II

In Phase II, the Evaluation team collected data through interviews conducted with TSA headquarters and field personnel and on-site observations made at federal and private pilot airports. The team performed analyses of the data to draw conclusions regarding private contractor and federal screening operations performance. Phase II activities included:

- The collection of performance measures and data from TSA headquarters and the field.
- Airport visits to develop an understanding of the operating environments at both private pilot and federal airports and identify systematic biases in the data collection.

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



- A stakeholder focus group whereby external stakeholders could review the measures, methodology, and comparison sites to be used in the Evaluation. Membership of the focus group is described in the Screening Operations Performance Evaluation Plan (Phase I Report).
- The examination and cleansing of data collected by the team to determine the results of outcome variables. The team compared the performance results found at the private pilot airports to those found at federal airports within the same size and risk category. Through statistical tests, the team was able to draw inferences and conclusions.

Results from the Phase II analysis are documented in this report.

### 4.2 EVALUATION DESIGN

TSA identified effectiveness and efficiency as the two main areas in which it sought to compare the screening operations of private pilot airports to those of federal airports. For the purpose of this evaluation, the team defined effectiveness as a screening operation's ability to detect and prevent prohibited items from entering an airport's sterile area. The team labeled this criterion "security effectiveness," and used measurements including covert testing, TIP data, gate screening and re-certification testing. The team defined efficiency as the cost incurred by a security screening operation and labeled this criterion "cost to government." The Evaluation team also added a third criterion to compare the performance of the two screening models. The team believed there was a need to account for the impact that the screening operation has on its customers (passengers) and stakeholders (air carriers and airport management). It labeled this criterion "Customer/Stakeholder Impact." A more comprehensive discussion of the rationale behind the performance criteria selection can be found in the Screening Operations Performance Evaluation Plan (Phase I Report).

In order to design the Evaluation, the team sought to identify performance measures used by both the private contractor and federal screening models. In doing so, the team first examined the program requirements set forth in ATSA. Most notably, ATSA did not establish the Private Pilot Program in a manner that was optimal for evaluating screening operations performance. Under ATSA, "participation in the pilot program is limited to no more than one airport from each of the five airport risk categories."<sup>4</sup> This limitation of five airports, one per category, restricted the Evaluation. It would have been preferable to assess the different screening models through a single design, such as conducting federal and privatized operations simultaneously in the same airport or randomly assigning large numbers of airports to federal or private contractor conditions. The Private Pilot Program, as defined by ATSA, did not allow for sound, controlled comparisons to be made between the private contractor and federal screening models. Hence, the Evaluation team designed the most feasible analytical approach given this limitation.

When designing the Evaluation, the team also considered using a number of additional performance measures; however, two main factors prevented their use. First, conditions across private pilot airports were not controlled in a uniform manner and could not be assumed to be consistent. For example, TSA granted the screening companies operating at the five private pilot airports latitude in the recruitment of screening personnel, the supervision and scheduling of screeners, and the management of materials and overhead. As a result, these latitudes were employed differently at each of the private pilot airports. Factors such as the extent of FSD oversight at each airport, FSD staff and screener staffing levels, and the lack of clearly defined roles and responsibilities for the contractor also caused variations within the five private pilot airports, making each private pilot airport unique in many regards. Second, a sufficient

---

<sup>4</sup> Public Law 107-71 - Aviation and Transportation Security Act, Section 108.

amount of historical data did not exist for the evaluation of security screening operations performance at the private pilot airports. This deficiency was due to various non-standardized screening models that were in place prior to the arrival of the current screening contractors. Historical data therefore did not represent performance of the screening functions as they are currently designed.

### 4.3 EVALUATION METHODOLOGY

After identifying sources for performance data, the Evaluation team followed a systematic approach to reduce the universe of available data to a set of usable measures. The first step was to identify existing data related to screening operations. The team accomplished this step through meetings and interviews with TSA headquarters personnel, airport personnel, and external stakeholders. Next the team developed a logic model, which is a tool used to describe the causal relationships between various measures and their outcomes. The team then reviewed the existing data to see if it fit into the logic model as an input, process, or output measure. Only those measures related to the outcome were considered for the Evaluation. Finally, the team asked a series of questions regarding these candidate measures to determine their known limitations and suitability for the Evaluation. These questions included:

- What is the availability of the data (e.g., is the data being collected or could it be collected)?
- What is the quality of the data (e.g., is the data collected in a consistent manner)?
- What is the quantity of the data (e.g., is there a sufficient amount of data collected)?
- Do systematic biases exist (e.g., is the data measured only under certain conditions)?

The result of this process was a set of performance measures for use in the security effectiveness and customer/stakeholder impact comparisons.

In addition, prior to selecting performance measures, the Evaluation team had to consider another important factor: the time frame available to conduct the Evaluation. TSA requested that the team complete the Evaluation in ninety days. This limited timeframe required the team to rely on data already collected by TSA at both federally and privately screened airports, as opposed to identifying and collecting new performance measures beyond those which already existed. With this constraint in mind, the team identified the measures needed to evaluate screening performance in the areas of security effectiveness, customer/stakeholder impact, and cost to government.

## 5.0 CONCLUSIONS

This section provides the overall findings from the comparison of the five privately screened airports to federally screened airports in the areas of security effectiveness, cost to government, and customer and stakeholder impact. The purpose of the Evaluation was to determine whether the level of screening services provided by private screening operations was equal to or greater than that provided by the federal government. A finding that a particular airport outperformed its counterparts does not imply that it outperformed all airports within its category, but it does imply the airport performed above the average level of the federal airports to which it was compared.

The section includes a general discussion of the conclusions by evaluation criteria area.

### 5.1 SECURITY EFFECTIVENESS CRITERIA

This subsection discusses the results of each measure in the area of security effectiveness. These measures included covert testing, TIP data, gate screening, and re-certification testing. An overall conclusion is provided at the end of this subsection.

#### Covert Tests

Kansas City performed significantly better on TSA covert testing than other airports in its category. Apart from this, there were no statistically significant differences between private pilot airports and federally screened airports in performance on TSA covert tests or in covert tests administered by DHS and GAO.

#### TIP Data

There was no evidence to suggest that TIP performance varied between privately screened and federally screened airports.

#### Gate Screening

At the two airports at which gate screening data was available, Kansas City and Rochester, results suggested that these airports outperformed their federal counterparts in the detection of threat and prohibited items. However, due to the sporadic practice of gate screening, the team's level of confidence in these findings is not high.

#### Re-certification Tests

The analysis of the Standardized Proficiency Review (SPR), or job knowledge, re-certification test results indicates that San Francisco outperformed the federally screened airports within its category. The test results also indicate that all other privately screened airports performed the same as the federally screened airports within their respective categories.

The analysis of the Image Proficiency Review (IPR) re-certification test results indicates Tupelo outperformed the federally screened airports in its category. All other privately screened airports performed the same as the federally screened airports within their respective categories.

The analysis of Checked Baggage re-certification test results indicates that San Francisco, Kansas City, and Rochester performed at the same level as federally screened airports within their respective categories. Data was not available for Jackson Hole and Tupelo.

### **Security Effectiveness Conclusion**

Overall, Kansas City outperformed federally screened airports in its category on security effectiveness. There is no evidence that security effectiveness in the other four pilot program airports was materially different from security effectiveness in federally screened airports.

## **5.2 CUSTOMER SATISFACTION AND STAKEHOLDER IMPACT CRITERIA**

This subsection discusses the results of each measure in the area of customer satisfaction and stakeholder impact. These measures included overall customer satisfaction results from the Customer Satisfaction Index survey, complaints data from the TSA Contact Center and at the airport level, and passenger wait times. An overall conclusion is provided at the end of this subsection.

### **Overall Customer Satisfaction**

The Customer Satisfaction Index included questions on overall customer satisfaction, screener thoroughness of screening, and passenger confidence in the security process. When Kansas City and San Francisco overall customer satisfaction survey results were compared to the federally screened airports' survey results, Kansas City under-performed and San Francisco outperformed the federally screened airports within their respective categories. Kansas City outperformed federally screened airports in its category on screener thoroughness of screening and under-performed on passenger confidence in the security process. San Francisco performed the same as federally screened airports in its category on screener thoroughness of screening and under-performed on passenger confidence in the screening process. Conclusions resulting from this comparison are limited due to the distribution of the Customer Satisfaction Index survey to a small number of airport locations. A significant number of Category II, III and IV airports were not included in the distribution.

### **Complaints**

Two data sets were used to determine levels of customer satisfaction: customer complaints received at airports and customer complaints received at the TSA Contact Center. Results from analyses of both data sets showed no significant differences between privately screened airports and federally screened airports in the same categories.

### **Wait Time**

Passenger checkpoint wait time analyses for San Francisco and Kansas City show that San Francisco and Kansas City had significantly shorter wait times than federally screened airports within their respective categories. Sufficient wait time data were not available for the private pilot and federal Category II, III, and IV airports. Therefore, an analysis of these airports was not completed.

### **Stakeholder Impact**

Findings from the qualitative stakeholder analysis include the following:

- Stakeholders experience little to no differences between private and federal screening operations.
- Stakeholders were satisfied with local TSA staff performance but critical of TSA headquarters.

### **Customer Satisfaction Conclusion**

San Francisco and Kansas City customer satisfaction analysis consists of mixed results when compared to federally screened airports within their respective categories. Customer satisfaction data for federally screened Category II, III, and IV airports was not available. Therefore, no comparison could be made.

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



### 5.3 COST

This subsection describes the conclusions made regarding cost to the government for private screening operations. As stated above, costs for the five privately screened airports were not significantly different from the estimated cost of a federal screening operation at that same airport.

Tables 5-1 and 5-1a illustrate the estimated costs for privately contracted screening and TSA screening for each pilot airport. They show the difference between the two costs in dollars and as a percentage of private costs based on two separate assumptions discussed below.

The first assumption states that the number of screeners, or staffing levels, for each of the five pilot airports would be based on predictive measures that take into account the staffing levels at other similar airports. Since these results are based upon a staffing prediction resulting from a regression model, the range of prediction error is presented in Table 5-1. Cost differences within the range of error may arise from chance variation in staffing levels, and therefore, the Evaluation team cannot conclude that these differences are meaningful. The second assumption states that the staffing level used by the federal government at each of the pilot airports would be the same as the number of screeners currently used by the contractor.

Table 5-1 provides cost data based on the assumption that staffing levels at each of the pilot airports would be determined using predictive measures that were derived from staffing levels at other similar airports.

**Table 5-1 Difference in Estimated Cost of Contracting and TSA Screening at Pilot Airports, Assuming That TSA Would Have Staffed According to Prediction Based on Experience of Other Airports (in \$ Millions, December 1, 2002 – November 30, 2003)**

	Est. Cost of Contracted Screening	Est. Cost of TSA Screening	Estimated Difference Compared to Contracted Screening		Range of Error in Estimate*
			Amount	Percent	Amount
San Francisco	\$118.9	\$119.0	\$0.15	0.1%	± \$20.7
Kansas City	51.1	59.0	7.9	15.5%	± 9.1
Rochester	15.3	14.2	-1.2	-7.6%	± 2.4
Jackson Hole	3.5	3.6	0.1	2.3%	± 1.5
Tupelo	1.5	1.1	-0.4	-28.4%	± 0.540

\* 95% significance level.

Sources: Computed and/or estimated using expenditure data provided by TSA.

For San Francisco International Airport, under the assumption that the staffing level would be determined using predictive measures that were derived from staffing levels at similar airports, the private screening operation is estimated to be less costly than a federal screening operation---\$118.9 million for private contractor screening as opposed to \$119 million for federally operated screening. Results are well within the range of possible prediction error, so the Evaluation team cannot conclude that these differences are meaningful.

## Screening Operations Performance Evaluation

Phase II: Evaluation Report – *Summary*

April 16, 2004



For Kansas City International Airport, under the assumption that the staffing level would be determined using predictive measures that were derived from staffing levels at similar airports, the private screening operation is estimated to cost less than a federal screening operation at that airport---\$51.1 million as opposed to \$59 million. The \$7.9 million difference is within the range of prediction error. The Evaluation team cannot, therefore, conclude confidently that federal screening operations would have been more costly.

For Greater Rochester International Airport, the contractor maintained a higher average number of screeners on staff throughout the year than the reduced federal numbers called for in the targeted positions issued in June 2003. Therefore, the predicted staffing levels for TSA screening were close to what would have been expected if the targeted positions had been filled. Under the assumption that TSA would have staffed according to the prediction based on the experience of other airports, the private screening operation is estimated to be more costly than a federal screening operation---\$15.3 million as opposed to an estimated \$14.2 million---this is also within the range of prediction error, so the team cannot conclude that this difference is meaningful.

For Jackson Hole Airport, under the assumption that TSA would have staffed according to the predictive measures based on the experience of other airports, the private contractor screening operation was estimated to be less costly than the federalized screening operation---\$3.5 million as opposed to \$3.6 million. This difference is within the range of prediction error and cannot be interpreted as meaningful.

For Tupelo Regional Airport, under the assumption that TSA would have staffed according to the prediction measures based on the experience of other airports, the private contractor screening operation was estimated to be more costly than the federalized screening operation---\$1.5 million as opposed to \$1.1 million. However, staffing patterns at these federalized airports varied considerably from airport to airport, and small changes in numbers of staff at these airports make predictions subject to a wide range of error. This difference falls within the range of prediction error.

Table 5-1a provides cost data based on the assumption that staffing levels at each of the pilot airports would be determined by using the same number of screeners used by the current contractors at those airports.

**Table 5-1a Difference in Estimated Cost of Contracting and TSA Screening at Pilot Airports, Assuming That TSA Would Have Staffed With Same Number of Screeners as Contractors (in \$ Millions, December 1, 2002 – November 30, 2003)**

	Est. Cost of Contracted Screening	Est. Cost of TSA Screening	Estimated Difference Compared to Contracted Screening	
			Amount	Percent
San Francisco	\$118.9	\$116.9	-\$1.9	-1.6%
Kansas City	51.1	52.5	1.4	2.8%
Rochester	15.3	16.2	0.8	5.5%
Jackson Hole	3.5	3.8	0.2	6.5%
Tupelo	1.5	1.3	-0.2	-10.3%

*Sources: Computed and/or estimated using expenditure data provided by TSA.*

Under the assumption that TSA would have staffed with the same number of screeners as the contractors, the private contractor screening operations in San Francisco is estimated to be more costly than the federalized screening operation -- \$118.9 million as opposed to \$116.9. At Kansas City International Airport, the private screening operation is estimated to be less costly than a federal screening operation—\$51.1 million as opposed to an estimated \$52.5 million. At Rochester International Airport, private screening operations is estimated to be less than the federally operated screening--- \$15.3 million as opposed to \$16.2 million at Rochester International Airport. At Jackson Hole, private contractor screening is estimated to be less costly than the federally operated screening---\$3.5 million as opposed to \$3.8 million. Finally, at Tupelo Regional Airport, the private contractor screening is estimated to be more costly than the federally operated screening -- \$1.5 million as opposed to \$1.3 million. This is largely because the contractor operated with more screening staff than predicted for a federalized operation using information about staffing patterns at other Category IV airports.

#### **5.4 USES AND LIMITATIONS OF THE STUDY**

This study concluded that the privately screened airports appear to have met the standard articulated in ATSA that contract screening providers perform at the same level or better than federally screened airports. There is no evidence that these airports are not meeting the ATSA requirement.

However, the ability to generalize these conclusions to any future expansion of privately screened airports is limited by the non-random selection of the participant airports, the small sample size, the lack of a steady state environment, and the limited data available. The five-airport, non-random sample limits the precision of the study's findings. The data collection systems available are in many cases immature and were not designed to capture federal/private differences. Additionally, the number of observations for each measure was limited.

Other factors that may limit the number of differences in the findings are the “degrees of freedom” discussed in Section 3.2. The study does not conclude that opening these degrees of freedom will lead to greater observed differences between privately screened and federally screened airports. However, allowing private contractors greater latitude in these areas may lead to differing levels of performance in each of the three criteria tested.

TSA may want to address these issues in any future decisions regarding private screening operations. If TSA desires a more robust comparison of private screening operations to federal screening, it should consider three steps: open some of the degrees of freedom in a controlled manner; provide a larger, well-designed sample of airports; and improve its data collection systems.